



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/676,377	09/30/2003	Ulrich Neumann	06666/156001/USC-3345	3241
20985	7590	05/04/2005	EXAMINER	
FISH & RICHARDSON, PC 12390 EL CAMINO REAL SAN DIEGO, CA 92130-2081			WOODS, ERIC V	
			ART UNIT	PAPER NUMBER
			2672	

DATE MAILED: 05/04/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

<b>Office Action Summary</b>	<b>Application No.</b> 10/676,377	<b>Applicant(s)</b> NEUMANN ET AL.	
	<b>Examiner</b> Eric V Woods	<b>Art Unit</b> 2672	

**-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --**

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

- 1) ☒ Responsive to communication(s) filed on 30 September 2003.
- 2a) ☐ This action is **FINAL**.                      2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

- 4) ☒ Claim(s) 1-44 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-44 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_\_ are subject to restriction and/or election requirement.

**Application Papers**

- 9) ☒ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 30 September 2003 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All    b) ☐ Some \*    c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_\_.
3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

- |   |   |
|---|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)   | 4) <input type="checkbox"/> Interview Summary (PTO-413)<br>Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)  | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)             |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)<br>Paper No(s)/Mail Date <u>20050329, 20050304</u> . | 6) <input type="checkbox"/> Other: _____  |

## **DETAILED ACTION**

### ***Priority***

1. Priority claim for CIP is noted and given, but applicant is reminded that the filing date of the parent application is **after** the filing date of the provisional and before its expiration; as such, this application will receive priority for the provisional filing date, which is earlier, regardless of the status as a CIP.

### ***Specification***

2. The lengthy specification has not been checked to the extent necessary to determine the presence of all possible minor errors. Applicant's cooperation is requested in correcting any errors of which applicant may become aware in the specification.

3. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed.

4. The following title is suggested: "Augmented Virtual Environments With Projection of Real-time Video."

5. The Background section of the specification is objected to for the following reasons:

### **Content of Specification**

(e) Background of the Invention: See MPEP § 608.01(c). The specification should set forth the Background of the Invention in two parts:

(1) Field of the Invention: A statement of the field of art to which the invention pertains. This statement may include a paraphrasing of the applicable U.S. patent classification definitions of the subject

matter of the claimed invention. This item may also be titled "Technical Field."

- (2) Description of the Related Art including information disclosed under 37 CFR 1.97 and 37 CFR 1.98: A description of the related art known to the applicant and including, if applicable, references to specific related art and problems involved in the prior art which are solved by the applicant's invention. This item may also be titled "Background Art."

Specifically, no section regarding Prior Art or Related Art is included, and no discussion of information disclosed under 37 CFR 1.97 and 1.98. Applicant is required to amend the specification to include at least a discussion of the most relevant pieces of prior art. Also, no mention of the Field of the Invention is made.

6. The Summary section is objected to because it does include at least some indication of improvements of the present invention over the prior art, and further merely recites sample claims. The content of the summary section should be:

**Content of Specification**

- (f) Brief Summary of the Invention: See MPEP § 608.01(d). A brief summary or general statement of the invention as set forth in 37 CFR 1.73. The summary is separate and distinct from the abstract and is directed toward the invention rather than the disclosure as a whole. The summary may point out the advantages of the invention or how it solves problems previously existent in the prior art (and preferably indicated in the Background of the Invention). In chemical cases it should point out in general terms the utility of the invention. If possible, the nature and gist of the invention or the inventive concept should be set forth. Objects of the invention should be treated briefly and only to the extent that they contribute to an understanding of the invention.

7. The disclosure is objected to because of the following informalities:

-Page 5, paragraph [0020], the terms "110, 120, 130" are stated, wherein it should read "110, 120, and 130";

-Page 6, paragraph [0022], bottom of the page, the sentence reads, "can provide a rapid an accurate" where it should read "can provide a rapid and accurate".

Appropriate correction is required.

### ***Drawings***

8. Examiner accepts the drawings.

### ***Claim Rejections - 35 USC § 112***

9. The following is a quotation of the first paragraph of 35 U.S.C. 112:

The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

10. Claims 1-44 are rejected under 35 U.S.C. 112, first paragraph, because the specification, while being enabling for generating three dimensional models from range sensors, projecting real-time video, and visualization, does not reasonably provide enablement for "at least one image sensor". The specification does not enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and/or use the invention commensurate in scope with these claims.

Specifically, the specification states on pages 24-25, paragraphs [0064] and [0065], that a stereo camera is necessary and this is illustrated in Fig. 2 for a mobile image sensor. Also, even for fixed placements, it appears to be necessary for the camera to be a "stereo camera" based on the specification. Therefore, independent claims 1, 14, 24, 32, 40, and 42 all clearly fail to be enabled

11. The following is a quotation of the second paragraph of 35 U.S.C. 112:

The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter, which the applicant regards as his invention.

12. Claims 1-44 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

13. Specifically, the metes and bounds of claims 1-44 are rendered unclear because the specification states on pages 24-25, paragraphs [0064] and [0065], that a stereo camera is necessary and this is illustrated in Fig. 2 for a mobile image sensor. Also, even for fixed placements, it appears to be necessary for the camera to be a "stereo camera" based on the specification. Therefore, independent claims 1, 14, 24, 32, 40, and 42 all clearly fail to define the word 'camera' in the sense that it is understood in the art, in this case, as being a stereo camera. If applicant is claiming only a single camera system (or the claim encompasses such devices) – then applicant is asked to clarify what this term means.

14. The term "morphological" in claims 28 and 36 is a relative term, which renders the claim indefinite. The term "morphological" is not defined by the claim, the specification does not provide a standard for ascertaining the requisite degree, and one of ordinary skill in the art would not be reasonably apprised of the scope of the invention.

15. Claims 28 and 36 are rejected under 35 U.S.C. 112, second paragraph, because the term "morphological" is not defined and has no art-accepted meaning. The term "morphological" as defined by most dictionaries is taken to mean, "having to do with

form or structure of a system or a biological organism". Applicant must provide a definition for this term and amend the claim accordingly. As stated above, the specification does not provide any standard for understanding this term.

***Claim Rejections - 35 USC § 103***

16. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

17. Claims 1-2, 5, 7-9, 11, 14, 21, and 40 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kumar et al (US PGPub 2001/0038718 A1)('Kumar') in view of Haala et al (Haala, N. and C. Brenner. "Generation of 3D City Models From Airborne Laser Scanning Data.")('Haala') (Claims 1, 14, and 40 are the same thing – system, method, and computer program product that all perform the same tasks, so the rejection is equally valid on all of them without further comment).

18. As to claims 1, 14, and 40,

A method comprising:

-Generating a three dimensional model of an environment from range sensor information representing a height field for the environment; (Haala section 1, pgs. 105-106, see especially Fig. 1 where part of a 3-D model derived from an airborne LIDAR range sensor is shown, where clearly this constitutes a "height field" for the urban environment)(Kumar Fig. 1, wherein the video imagery sent to the computer terminal clearly shows the three-dimensional coordinates (100) of the portion of the image the

Art Unit: 2672

operator is looking, wherein paragraphs [0038-0041] teach the overlay of video imagery onto a 3-D map from aerial sources, and [0052-0054] teach the alignment of video with 3D coordinates with DTM information (e.g. see Fig. 4 with step 406 showing this check))

-Tracking orientation information of at least one image sensor in the environment with respect to the three-dimensional model in real-time; (image sensor in the environment, for example Kumar Figs. 1 and 2, element 102, the platform is taught to provide

'engineering support data', e.g. GPS/INS, height, etc., in [0038-0041]. Specifically, it is disclosed that the system of Kumar can take any camera on a moving platform and correlate the imagery to previously captured imagery stored in the database in [0006];

and that it is well-known in the art to use mobile coordinate information within some fixed world coordinates (e.g. latitude and longitude), where the platform transmits "engineering support data".)

-Projecting real-time video imagery information from the at least one image sensor onto the three dimensional model based on the tracked orientation information; and (Kumar [0034-0037] discloses the use of projective techniques for aligning video frames and details of how this technique is used; [0052] clearly teaches that the video imagery is combined into a single 3D projective view using fine alignment block 222 as shown in Fig. 2 for example, and specific algorithms are taught in [0067] and [0069])

-Visualizing the three-dimensional model with the projected real-time video imagery.

(Kumar [0098] and [0108] teaches that the system may be used for real-time

applications and that the system can be modified to function in a real time environment,



with the operator view shown in Fig. 1 with the representations derived from the Haala system)

Kumar teaches a system whereby an existing three-dimensional environment map is created and then a mobile platform (an aerial platform is shown as an exemplar, but the techniques are certainly not limited to it) having a camera and position- and attitude-determining equipment that provide "engineering support data" such that the video that it records and sends back to the user's terminal can be overlaid onto the three-dimensional model using projective transformations. The Kumar reference incorporates GIS systems so that terrain information and similar can be integrated with video mosaics (see [0008] for example). However, Kumar does not teach the specific use of a height field for an environment acquired from a range sensor, although given that the system of Kumar does integrate data from GIS and geospatial databases that is known to contain terrain data [0008] and [0069], intrinsically that information would have to be acquired in some manner.

The Haala reference teaches the use of an airborne LIDAR platform that obtains three-dimensional height data and then overlays it with information from a 2-D GIS to form a three-dimensional representation of urban terrain. This system is a perfect example of sensor fusion to create the three-dimensional system database or geospatial database of Kumar that is required. The references are directed to the same problem solving area, as both require and teach methods of generating three-dimensional terrain maps with video imagery, with Kumar focusing on doing so from mosaics but also taking as input DTM (such as that provided by USGS) data (that is,

Art Unit: 2672

digital terrain models that have known height data and similar), in a manner such as that cited by Haala (p. 106, section 1, wherein the 2D GIS tools are cited and how they are overlaid is discussed). This capability of Haala would allow the effective overlay of the video mosaic data of Kumar onto the range data combined with GIS of Haala to obtain applicant's invention. It would have been obvious to combine the system of Haala with Kumar because Haala will provide increased registration accuracy (based on integration with existing GIS systems) and thusly will allow more effective interleaving with the video input of Kumar to get more accurate projective transformations.

19. As to claim 2, Haala teaches in section 2 (page 106) that the generation of the three-dimensional model relies on the extraction of geometric primitives that correspond to items like a roof of a building, and this data is of course from the range sensor.

Specifically, the first sentence of section 2.1 (page 107) specifically states, "range image segmentation, which aims at dividing the object surface into patches that can be described parametrically ..." The work of Haala does not teach away from this, merely stating in the next line that their work is limited to a subset of that because that is all that is necessary, with multiple algorithms tested for use in their system. Examiner further maintains that it would have been obvious to use a parametric function to fit the geometric primitives, with backing provided by Kumar ([0008], [0034], [0044-0048]), who uses parametric fitting to fit the images to reference images, wherein those techniques are known to be used in fitting three-dimensional range data as Haala teaches. Finally, see Kumar [0108] where it is taught that a three-dimensional mesh, e.g. using geometric primitives, is parametrically mapped onto the image mosaic with known

information on height, wherein such a technique could easily be used on the range data of Haala. Motivation and combination is incorporated in its entirety from the rejection of the parent claim.

20. As to claim 5, the Haala reference clearly teaches different levels of detail in Figs. 2 and 3, and Fig. 1 clearly shows the grid used to hold the data after it is captured (pgs. 106-107). Specifically, in the conclusion on pg. 111 teaches that varying the level of detail is known in the art, e.g. "The level of detail, which can be reached for a reconstitution has to be examined depending on the density of measured laser points."

Based on these statements, it is obvious that the points are projected onto a regular grid in any case (based on rejection to claim 1 above, where projection is taught) – the case of user-defined resolution is as taught above; the reference teaches that it would have been obvious to modify the reference to use multiple resolutions or levels of detail.

Motivation and combination are taken from the parent claim.

21. As to claim 7, Haala is a LIDAR system (specifically, see pg. 106, section 1, or the tagline on Fig. 1, "DSM measured by airborne laser scanning." Motivation and combination are incorporated by reference from the parent claim.

22. As to claim 8, Kumar clearly teaches the use of visual comparison to determine the location of a viewed image with respect to reference image(s) in [0008-0009] with emphasis on [0032-0034] where the technical details of the process is disclosed.

23. As to claim 9, Kumar in [0041] teaches that the "engineering support data" provided by the camera platform (as mentioned in [0032—0034]) includes GPS (global positioning system, which is prima facie a satellite navigation system), an INS (inertial

Art Unit: 2672

navigation sensor) and obviously the camera input (visual input) as recited in the rejection to claim 8; with both claims, only the primary reference is utilized, no separate motivation or combination is required and that from the rejection to the parent claim is herein incorporated by reference.

24. As to claim 21, it is a duplicate of claim 8 with the additional limitations of claim 9. As such, the rejections to both claims immediately above are incorporated herein by reference, along with their motivation and combination.

25. As to claim 11, Kumar teaches in [0052-0054] the use of depth maps with video images; wherein the system computes depth using a reference image and acquired video imagery, with the technical details spelled out in [0063-0064]. Further, Kumar teaches the use of algorithms designed to minimize the occlusion effect [0009, 0078], which utilize hidden surface removal, thusly performing the recited limitation.

26. Claims 3, 18, and 41 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kumar in view of Haala as applied to claim 2 above, and further in view of You et al (You, S. and U. Neumann: "Automatic Object Modeling for 3D Virtual Environment." 1998)('You')

27. As to claim 3, references Kumar and Haala do not expressly teach this limitation, except that reference Haala clearly teaches the identification of structures in the range data, for examples Figs. 3-4 and 6-7 where such structures are fitted to range data. Reference You clearly teaches on page 22 the use of superquadratics in the use of fitting geometric primitives using parametric fitting for objects such that the geometric shape description of a 3D object can be obtained (see for example page 23, upper half

Art Unit: 2672

of the pages, still part of section 1). Motivation and combination is provided by the fact that superquadratics provide more accurate modeling than the planar surface technique of Haala while being capable of deformed to represent many irregular curved objects (pg. 23), thusly providing the motivation to combine You with Kumar and Haala, and the motivation to combine Kumar and Haala in the first place is incorporated by reference from the parent claim.

28. As to claims 18 and 41, the rejections to claims 2 and 3 above are herein incorporated by reference. This claim is essentially the same as that of claim 3, with the additional limitation of identifying structures in the range sensor information, which reference Haala very clearly teaches in Figs. 1 and 3, wherein the system of Haala function by segmenting planar regions – see section 1 on pages 105-106 – to identify and generate buildings. Motivation and combination is taken from claim 3 and incorporated herein by reference. Claim 41 is merely software implementing this.

29. Claims 4, 10, 15-17, and 20 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kumar in view of Haala as applied to claim 2 above, and further in view of Arpa et al (US 2003/0085992 A1).

30. As to claims 4, 15, and 20, references Kumar and Haala do not expressly teach this limitation. Reference Arpa teaches this limitation, specifically in Fig. 1 where multiple cameras (108, 110, 114) are shown being linked to an image processor, where the system in Fig. 2 then takes all the video inputs and projects them onto a 3D model using 3D model generator 210; the results are shown in for examples Figs. 7 and 8 where an object that is moving is tracked (see [0026-0028]). Clearly, the system of

Art Unit: 2672

Arpa projects the results of multiple video cameras onto a known three-dimensional model of a scene. The specific limitation of "refining the three-dimensional model based on object surfaces" is clearly performed by Haala, as for example in Figs. 10-13, where initial reconstruction of the scene is shown and the adjusted or corrected version is shown (which is comparable to the recited "further ... refining" recited in the claim).

Clearly, it would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the systems of Kumar and Haala with that of Arpa, as the system of Kumar teaches the use of one camera; the system of Arpa would prima facie enhance the capability of that system by allowing it handle multiple cameras (Kumar suggests multiple cameras in [0020-0023, 0040] where it is taught that multiple reference images can be used to eliminate parallax, which to one of ordinary skill in the art is known to require two or more cameras, and further multiple "reference images" are taught in [0045], which would obviously encompass multiple images of the same scene taken from different angles, e.g. the claimed multiple sensors.) (Note that claims 15 and 20 are very similar to / obvious variants of one another and are covered by the above rejection).

31. As to claim 10, the system of Arpa clearly covers the scenario of both multiple image sensors as shown above, and claim 1 itself teaches the application of real-time video imagery information. The limitation of multiple video streams being projected onto the three-dimensional variant is taught by Arpa, where all the views of the cameras are incorporated into the 3D model so that a user can choose between desired views with

respect to the structural model, for example Figs. 1 and 2. Motivation and combination are taken from the rejection to claim 4 above.

32. As to claim 15, it is a substantial duplicate of claim 4; the rejection is the same.

33. As to claim 16, Kumar teaches in [0036] the use of that system to provide "rewind" and "replay" functionality and that such video can be stored on videotape or digitally to be viewed at a later point.

34. As to claim 17, Arpa teaches in Fig. 8 the use of a 'synthetic' view to show the location of a moving object, which is equivalent to the recited limitation of computing the view of the virtual camera from a viewpoint not that of one of the image sensor; see [0047] specifically for the details on how the virtual camera works. Motivation and combination is taken from the rejection to claim 15 and incorporated by reference.

35. Claims 6 and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kumar in view of Haala as applied to claim 1 above, and further in view of Weinhaus et al (Weinhaus, F. and V. Devarajan. "Texture Mapping 3D Models of Real-World Scenes.")

36. As to claim 6, references Kumar and Haala do not expressly teach this limitation, except that Kumar teaches the use of a triangulated mesh for reconstruction of 3D images and registration in [0104-0107], while reference Weinhaus teaches the limitations of tessellation and hole filling on pg. 349 where it is stated that terrain elevation tiles are tessellated into triangular meshes and on pg. 346 where methods using point-projection methods to fill holes. Therefore, it would have been obvious to modify the combination of Kumar and Haala to use the techniques recited in Weinhaus

Art Unit: 2672

to further refine and correct the three-dimensional models generated during the process, given that in pgs. 340-347 several different techniques are given for performing texture mapping on height-based data sets using projective techniques.

37. As to claim 19, it is substantially the same as claim 6, with the additional limitation of using a user-defined resolution is met in the rejection to claim 6 above, which uses only the Kumar and Haala references. Motivation and combination is taken from claim 6 above and incorporated in its entirety by reference.

38. Claims 13 and 23 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kumar in view of Haala as applied to claim 1 above, and further in view of Pryor et al (US 2004/0046736 A1)('Pryor').

39. As to claims 13 and 23, references Kumar and Haala do not expressly teach this limitation. Reference Pryor teaches the use of a virtual reality type user interface wherein large images are projected onto a screen, such as a wall screen [0043, 0056] and that many input devices, such as a head tracker are well known in the art and can be used with the invention of Pryor [0009, 0067, 0083-0085, particularly 0105 where it says that the user's head can be tracked for purposes of motion determination]. It would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the systems of Kumar and Haala with that of Pryor, given that in [0056] it is taught that using large screens makes it easier to interact with imagery and obviously this would facilitate data visualization and similar with the combination of Haala and Kumar.



40. Claims 24-28, 30-36, 38, and 42-44 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kumar in view of Arpa. Claims 24, 32, and 42 are all duplicates of each other, namely a method, system, and computer program product, and as such they stand or fall together.

41. As to claims 24, 32, and 42, the Kumar reference shows Fig. 1, wherein the video imagery sent to the computer terminal clearly shows the three-dimensional coordinates (100) of the portion of the image the operator is looking, wherein paragraphs [0038-0041] teach the overlay of video imagery onto a 3-D map from aerial sources, and [0052-0054] teach the alignment of video with 3D coordinates with DTM information (e.g. see Fig. 4 with step 406 showing this check)). Arpa clearly teaches a 3D model of the environment as in Fig. 2; with the 3D model generator 210.

Kumar teaches that the camera platform knows its position and can supply pose and position information, e.g. image sensor in the environment, for example Kumar Figs. 1 and 2, element 102, the platform is taught to provide 'engineering support data', based on e.g. GPS/INS, height, etc., in [0038-0041], e.g. position and orientation information in three dimensions. Kumar teaches that the video from the aerial platform as shown in Figs. 1 and 2 can be real time [0098]. Arpa clearly provides real-time functionality, since the desired system clearly is used for security monitoring [0006-0011], emphasis on [0012], and further the camera position and pose (e.g. orientation) of the system of Arpa are known prima facie [0030-0032] given that the technician who installs them marks their position on the three-dimensional model.

Next, the system of Arpa provides identification of a moving region in the background, as for example shown in Fig. 7 or in Figs. 11A-11D, wherein the formation of a background image using time averaging is well-known in the art and is also cited in [0041-0042], e.g. step 402 in Fig. 4, and wherein further details of the steps are outlined in [0043-0044], where the three-dimensional segmentation to find other non-moving objects is performed – specifically, another patent document is cited – EP1045591, which is the European version of US6618058 B1 – wherein that technique provides the recited functionality of generating a time-averaged version of the background (in this case, via histogram), which was incorporated into the Arpa document and is under assignment to the same entity. Clearly, this constitutes being “dynamically modeled from a time average of the real-time video imagery information”, as that US Patent teaches that it does time averaging, not static comparison (for example, 1:60-2:10, 3:3-4:67).

Arpa clearly places a surface that corresponds to the moving region in the three-dimensional model, e.g. the moving item shown in Fig. 7 is actually a silhouette that is known to obviously be a surface (albeit a three-dimensional one), but the system of Arpa shown in Fig. 4 clearly generates both 2D object silhouettes (406) and 3D ones (416) and outputs them. Therefore, both can be placed into the model (e.g. 2D in Fig. 7, 3D in Figs. 11A-11D), or in a synthesized view (Fig. 8)

Kumar [0034-0037] discloses the use of projective techniques for aligning video frames and details of how this technique is used; [0052] clearly teaches that the video imagery is combined into a single 3D projective view using fine alignment block 222 as

shown in Fig. 2 for example, and specific algorithms are taught in [0067] and [0069].

This clearly teaches "projecting real-time video imagery..." and the rest of that limitation, and Figs. 7, 8, and 11A-11D clearly show that such data is put into the model via projective techniques as recited by Kumar.

For visualization, Kumar [0098] and [0108] teaches that the system may be used for real-time applications and that the system can be modified to function in a real time environment, with the operator view shown in Fig. 1. Clearly, Arpa also teaches visualization in Figs. 7, 8, and 11A-11D.

Clearly, It would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the system of Kumar with that of Arpa, as the system of Kumar teaches the use of one camera; the system of Arpa would prima facie enhance the capability of that system by allowing it handle multiple cameras (Kumar suggests multiple cameras in [0020-0023, 0040] where it is taught that multiple reference images can be used to eliminate parallax, which to one of ordinary skill in the art is known to require two or more cameras, and further multiple "reference images" are taught in [0045], which would obviously encompass multiple images of the same scene taken from different angles, which would allow a security guard or user to more effectively visualize activity over an urban area for example.)

42. As to claims 25 and 33, see above, specifically where Arpa teaches that the system generates a two-dimensional silhouette as in step 406 in Fig. 4 and would be shown, e.g. Fig. 7 using the process of Fig. 5 (see [0045]).

43. As to claims 26, 34, and 43, the system of Arpa can generate a synthetic video view (e.g. Fig 8) wherein [0022, 0046-0048] the view is from any given viewpoint; that does not require a camera to be present there. In any case, clearly the system of Arpa performs the recited limitation. The optical center represents nothing other than the location of the camera; that is, the location that the (virtual) camera is focusing on from the frustum. Therefore, the system is capable of displaying real-time information in such a manner from a synthetic viewpoint, including indication of the movement of an object in two or three dimensions. As such, in order for the three-dimensional view to be accurate (e.g. Fig. 8), the viewpoint of the synthetic camera must be computed, which clearly involves drawing rays from the camera to the object in question to ensure that the object was on the same optical axis as the camera so that it would be visible. Clearly, such moving objects would be found in their own image plane per se, since an image plane would prima facie intersect with the optical axis of the synthetic view, which prima facie corresponds to the three-dimensional model.

Further, it is inherent and intrinsic that for a synthetic camera, such a moving region (e.g. that the position, orientation, and size) would be a function of the viewpoint, namely the intersection of the view lines from the camera through the three dimensional model, and further that, as shown in Fig. 7, the ground plane could be easily identified. Obviously, a three-dimensional model will have a reference or ground plane – it has to in order for a three-dimensional coordinate system to be adequately defined. Further, the size of the object will be set by the intersection of the optical axis and the image

plane of the object, thus automatically 'casting a ray' as recited in the claim. Motivation and combination are taken from claim 24 and incorporated herein by reference.

44. As to claims 27, 35, and 44, see the rejection to claim 24 above and further as taught in Arpa [0042] wherein it is stated the background is subtracted and foreground objects identified in real-time video imagery. For further example, see Fig. 4, wherein the background is computed (402) and then removed at step 404 and the foreground objects are output in step 406. Kumar also teaches the validation step when [0103-0112] he states that the reference imagery is compared to the real-time imagery and the three-dimensional model and key frames may be used for matching, or multiple sets of frames [0109-0112] which would inherently validate them, since if one frame was inadequate, then multiple frames would be necessary, which would mean that the one frame was found to be invalid, thusly the system would have to be validating the frames in the first place [0110]. Motivation and combination is taken from the parent claim.

45. As to claims 28 and 36, obviously Arpa establishes distinguishing foreground objects (step 406, Fig. 4, see also [0042-0044]). The Arpa reference states that it uses the method in US6618058 B1 as set forth in the rejection to claim 24. Further, that patent (Hanna) teaches in 2:1-25 the use of histograms for this application using threshold techniques (7:10-20) based on the histogram and it uses various types of filtering (2:1-25, 3:62- 4:10) that would qualify as morphological.

46. As to claims 30 and 38, Kumar clearly teaches that the image sensor platform provides information concerning position and orientation, e.g. from GPS systems and

Art Unit: 2672

inertial navigation systems and that is real-time, e.g. see the aerial platform cited in Figs. 1 and 2 and [0038-0041] where such data is termed 'engineering support data.'

47. Claims 31 and 39 are rejected under 35 U.S.C. 103(a) as being unpatentable over Arpa in view of Kumar as applied to claim 24 above, and further in view of Haala.

48. As to claims 31 and 39, references Arpa and Kumar do not expressly teach this limitation while Haala teaches it expressly -- Haala section 1, pgs. 105-106, see especially Fig. 1 where part of a 3-D model derived from an airborne LIDAR range sensor is shown, where clearly this constitutes a "height field" for the urban environment. Kumar teaches it implicitly in Fig. 1, wherein the video imagery sent to the computer terminal clearly shows the three-dimensional coordinates (100) of the portion of the image the operator is looking, wherein paragraphs [0038-0041] teach the overlay of video imagery onto a 3-D map from aerial sources, and [0052-0054] teach the alignment of video with 3D coordinates with DTM information (e.g. see Fig. 4 with step 406 showing this check).

Clearly, it would have been obvious to one having ordinary skill in the art at the time the invention was made to combine the systems of Kumar and Haala with that of Arpa, as the system of Kumar teaches the use of one camera; the system of Arpa would prima facie enhance the capability of that system by allowing it handle multiple cameras (Kumar suggests multiple cameras in [0020-0023, 0040] where it is taught that multiple reference images can be used to eliminate parallax, which to one of ordinary skill in the art is known to require two or more cameras, and further multiple "reference images" are taught in [0045], which would obviously encompass multiple images of the same scene taken from different angles, e.g. the claimed multiple sensors.)

***Allowable Subject Matter***

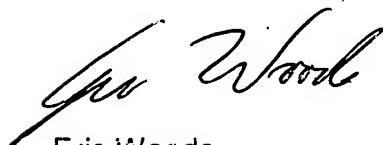
49. Claims 12, 22, 29, and 37 would be allowable if rewritten to overcome the rejection(s) under 35 U.S.C. 112, 2nd paragraph, set forth in this Office action and to include all of the limitations of the base claim and any intervening claims.

***Conclusion***

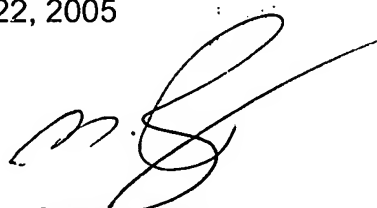
Any inquiry concerning this communication or earlier communications from the examiner should be directed to Eric V Woods whose telephone number is 571-272-7775. The examiner can normally be reached on M-F 7:30-5:00 alternate Fridays off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Michael Razavi can be reached on 571-272-7664. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

  
Eric Woods

March 22, 2005

  
MICHAEL RAZAVI  
SUPERVISORY PATENT EXAMINER  
TECHNOLOGY CENTER 2000